JOURNAL OF MATERIALS SCIENCE 22 (1987) 4036–4050

Thresholds and reversibility in brittle cracks: an atomistic surface force model

BRIAN R. LAWN, DAVID H. ROACH, ROBB M. THOMSON
Institute for Materials Science and Engineering, National Bureau of Standards, Gaithersburg, Maryland 20899, USA

A new picture of environmentally-enhanced fracture in highly brittle solids is presented. It is asserted that the fundamental relations for crack growth are uniquely expressible in terms of the surface force functions that govern the interactions between separating walls in an intrusive medium. These functions are the same, in principle, as those measured directly in the newest submolecular-precision microbalance devices. A fracture mechanics model, based on a modification of the Barenblatt cohesive zone concept, provides the necessary framework for formalizing this link between crack relations and surface force functions. The essence of the modification is the incorporation of an element of discreteness into the surface force function, to allow for geometrical constraints associated with the accommodation of intruding molecules at the crack walls. The model accounts naturally for the existence of zero-velocity thresholds; further, it explains observed shifts in these thresholds in cyclic load–unload–reload experiments, specifically the reduction in applied loading needed to propagate cracks through healed as compared to virgin interfaces. The threshold configurations emerge as thermodynamic equilibrium states, definable in terms of interfacial surface energies. Crack velocity data for cyclic loading in mica, fused silica and sapphire are presented in support of the model. Detailed considerations of the theoretical crack profiles in these three materials, with particular attention to the atomic structure of the "lattice" (elastic sphere approximation) at the interfaces, shows that intruding molecules must encounter significant diffusion barriers as they penetrate toward the tip region. It is concluded that such diffusion barriers control the fracture kinetics at low driving forces. At threshold the barriers become so large that the molecules can no longer penetrate to the tip region. This leads to a crucial prediction of our thesis, that the cohesive zone consists of two distinct parts: a "protected" primary zone adjacent to the tip, where intrinsic binding forces operate without influence from environmental influences; and a "reactive" secondary zone more remote from the tip, where extrinsic interactions with intruding chemical species are confined. The prevailing view of chemically enhanced brittle fracture, that crack velocity relations are determined by a concerted reaction with reactive species at a single line of crack-tip bonds, is seen as a limiting case of our model, operative at driving forces well above the threshold level. The new description offers the potential for using brittle fracture as a tool for investigating surface forces themselves.

1. Introduction

It has long been appreciated that cracks in brittle solids (i.e. solids with covalent/ionic bonding) can close and heal against a positive, non-zero driving force [1–3]; this although the crack walls may be exposed to chemically active species in the environment during their formation. Mica [4–8], silicate glasses [9–12] and sapphire [13] may be cited as "model" materials exhibiting this kind of behaviour. It is perhaps remarkable that while the thermodynamic roots of fracture reversibility (or irreversibility) have been long established, principally through the Griffith energy balance concept [1], there have been few systematic fracture mechanics studies of the phenomenon. In the present paper we address this shortcoming and make use of direct fracture reversibility observations to help develop a conceptually new picture of crack-interface processes.

Interest in the issue of crack reversibility has recently been revitalized by studies into the mechanisms of threshold behaviour in v–K or v–G (crack velocity against stress intensity factor or mechanical energy release rate) curves, specifically the tendency for cracks to approach a zero-velocity state at sufficiently low driving force [14–16]. Healing occurs at some point below this zero-velocity threshold. Investigations into such threshold behaviour do not abound in the scientific literature, due partly no doubt to time limits on data accumulation at "low" velocities (i.e. at $v \ll 10^{-9}\,\mathrm{m\,sec^{-1}}$). Perhaps the most detailed study, on soda–lime glass, is that of Michalske [17]. What has emerged from these limited studies is the conception

0022–2461/87 $03.00 + .12 © *1987 Chapman and Hall Ltd.*

of "crack blunting", a derivative of the classical Charles–Hillig model of stress-corrosion cracking [18], as the intrinsic cause of growth retardation. In the context of modern-day theories of brittle fracture, which picture the influence of environmental species in terms of a concerted chemical reaction with strained bonds at the tip of a sharp crack, such reversion to a blunting hypothesis suggests an ever-changeable tip structure. This picture clearly excludes any description of the brittle crack as a "fundamental lattice defect", in the sense, say, of a lattice dislocation with invariant Burgers vector [19].

Our present interest in this problem stems from work on soda–lime glass specimens containing aged indentation cracks [20, 21]. There it was noted that the cracks grow in the immediate post-indentation residual stress field, at first rapidly and then more slowly, ultimately (after about one day in water) coming to a stop. It was demonstrated that the post-indentation system had progressed down the v–K curve to the threshold configuration. Significantly, the strengths of specimens containing indentation cracks as dominant flaws showed no increase with further ageing time beyond this saturation point. According to the blunting hypothesis we may have expected such over-aged specimens to have been substantially stronger [20]. Our conclusion was that the cracks remained atomically sharp at all times, as if the intruding molecular species somehow do not have unrestricted access to the critical tip region (precluding, for instance, tip dissolution processes). It was thereby suggested that "internal closure forces", of the same kind as those responsible for healing, could account for the threshold phenomenon, without the need for recourse to any change in crack tip structure.

Our focus in the present study is accordingly on the nature of such internal forces, with special consideration as to how these forces may be incorporated into a fracture mechanics description. We shall develop a model based on the presence of cohesive interactions between the crack walls, i.e. in the crack region generally assumed to be completely free of all tractions. For the most part we comply with traditional continuum notions of crack geometry, regarding the environmentally interacting fracture system in terms of a linear elastic body in contact with a fluid reservoir. However, in setting up appropriate potential-function descriptions for the surface–surface interactions we shall introduce an important scaling element determined by the size of the intruding molecules. From this treatment thresholds emerge naturally as balance points between external driving and internal closing forces, in a manner strongly reminiscent of a Barenblatt equilibrium configuration [22]. A feature of the model is the determination of crack velocity and healing–repropagation thresholds as (separate) thermodynamic equilibrium configurations. Experimental observations of crack growth characteristics in mica, silicate glasses and sapphire (from both our own laboratories and literature sources) in load–unload–reload sequences in water-containing environments confirm the existence of the thresholds, but reveal strong kinetic effects during cyclic loading between these thresholds. Closer attention to the crack-wall geometry suggests that the kinetics in this region are associated with diffusion effects; at low K the crack becomes so narrow that the intervening species can no longer penetrate fully into the interactive cohesive zone near the tip. We explore this aspect in some detail, taking due account of the atomic structures of both the cracked material and the environmental fluid. In a companion paper [23] we shall propose detailed mathematical models based on these notions, but for present purposes a modified Barenblatt description will suffice.

As a result of our deliberations we shall paint a somewhat radical picture of fracture processes in the threshold region. The geometrical constraints referred to above are seen as a decisive factor in our description; we accordingly propose that the age-old representation of a brittle crack as a "continuum cavity" be replaced by that of a narrow interface, somewhat akin to a (dilated) grain boundary. The cohesion of this interface is uniquely determined by fundamental surface force functions, of the type measured directly on crossed mica cylinders in the newest microbalance devices [24]. Extraneous molecular species can play a vital role in determining equilibrium crack configurations by modifying (or augmenting) these surface force functions in the long-range, non-linear "tail" regions. This link with independent force measurements leads us to present fracture as a potentially useful alternative route to fundamental information on interatomic interactions: a route, moreover, that provides us with closer access to the short-range, primary cohesive region in the parent structure (by virtue of bond rupture in the near-tip region) and that extends the range of study to materials other than mica. The atomic structure of the crack interface also manifests itself as energy barriers to the intruding species, thereby accounting for the observed kinetic effects. The same barriers are expected to persist above threshold, with an attendant strong influence, via diffusion kinetics, on the entire v–G curve.

2. Surface force model of threshold behaviour

2.1. Existence of surface forces

The key to our modelling of threshold phenomena in brittle solids lies in the action of surface forces between opposing walls of the newly formed crack, even in the presence of contaminating species. We have already cited the observed tendency for cracks to heal spontaneously (if at reduced cohesive interfacial energy levels) in certain chemical environments. In addition, there is the wealth of independent experience from colloidal science [25], based on the well-respected "DLVO" theory [26, 27]. This theory, based strictly on continuum descriptions of matter, quantifies particle–particle surface interactions in solutions, most notably the issue of coagulation as against dispersion, in terms of a competition between attractive van der Waals and repulsive double-layer forces. Such forces, although perhaps weak (in relation to primary bonding forces), are long-ranged (in relation to atomic separations).

Over the last decade or so Israelachvili and

coworkers [28–32] have refined a technique for measuring surface forces directly. They use a delicate microbalance apparatus to determine the attractions (or repulsions) between two crossed cylinders of mica, down to atomic-scale separations. At reasonably large separations, say above 10 nm, their results essentially confirm classical DLVO theory. However, at small separations some significant departures from this theory are now being observed, revealing dramatic new features in interfacial interactions. Specifically, the surface–surface potential functions tend to increasingly large oscillations at near-contact, with a periodicity matching the diameter of the intervening molecules [33, 34]. Accordingly, the forces at a few molecular spacings can be much stronger than we might predict from the conventional descriptions. Here, then, is distinctive evidence for an element of atomic-scale discreteness in the fundamental interactions at narrow interfaces. This discreteness will be a recurring theme in our ensuing discussions.

To place the microbalance observations in the context of fracture mechanics we note that the crack opening displacements will necessarily span the entire range of the operative surface force function as we traverse the interface from tip to open mouth. In the spirit of Barenblatt we may conceive of a "cohesive zone" in the tip region, small in size in comparison with the crack length but large relative to atomic dimensions. Extraneous molecular species exert their strong influence on the fracture properties by significantly altering the energy states at the fully separated faces. Thus in terms of the Griffith–Barenblatt philosophy we should expect the sum effect of the interfacial interaction within the cohesive zone to be ultimately relatable to traditional surface energy parameters.

2.2. Fracture mechanics: equilibrium states

With this background we formulate a simple model for threshold behaviour, with the initial assumptions of linearity in elastic responses and continuity of matter. Suppose that the crack is driven by a remotely applied load as shown in Fig. 1. Opposing this applied load is a distribution of stresses, $p(x)$, over a cohesive zone, $0 \leqslant x \leqslant L$, behind the tip. Immediately we see the potential for describing the thresholds as quiescent balance points between external and internal forces. The $p(x)$ function includes any stresses arising from interactions with intruding chemical species. Ignoring for the moment the specific form this distribution function might assume (i.e. treating the crack tip as a "black box"), we may invoke the simple but powerful Griffith theorem to obtain, at equilibrium (designated by an asterisk) [19]

$$G^* = K_a^{*2}/E' = 2\gamma \tag{1}$$

where G is the mechanical energy release rate, K_a is the stress intensity factor associated with the applied loading, E' is Young's modulus (E for plane stress, $E/(1 - \nu^2)$ for plane strain, with ν Poisson's ratio), and γ is an appropriate surface energy. It is implicit in Equation 1 that the cohesive zone translates in an invariant manner with the crack tip, and that such

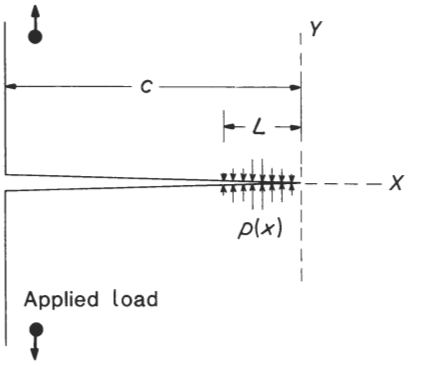

Figure 1 Schematic diagram of crack system, showing cohesive (Barenblatt) zone behind tip. Crack profile is defined by appropriate displacement field equations (e.g. Equations 4(a), (b)).

translations are reversible [34]. In this description γ represents the difference between the initial (unruptured) and final (ruptured) energy states at the crack plane. Following Orowan [35], we might immediately identify two experimentally relevant γ terms, one pertaining to the separation of the surfaces under inert conditions (or, more strictly, under conditions of solid–vapour equilibrium) and the other to separation in the pervasive presence of an interactive chemical environment. Thus the thresholds emerge naturally as null states of zero net crack driving force, $g = G - 2\gamma$ [2, 36]; for $g > 0$ the crack extends spontaneously, for $g < 0$ it contracts and heals spontaneously.

In terms of stress intensity factors, Equation 1 may be equivalently expressed in the form [14]

$$K_a^* = K_i \tag{2}$$

where K_i is an internal contribution associated with the cohesive stresses, $p(x)$. Again, $p(x)$ will be strongly modified by the presence of any interactive species. (In this work we define K_i as positive for $p(x)$ attractive, in accordance with our intuition that the cohesive stresses will usually act to resist crack advance.) Equation 2 is a statement of the Barenblatt condition for equilibrium [22], that the net stress intensity factor at the crack tip, $k = K_a - K_i$, should be zero; again, in this interpretation we predict the crack to run forward or backward spontaneously for k greater or lesser than zero, respectively. In general, we can express K_i as an integral of the closure stresses over the cohesive zone,

$$K_i = \int_0^L F(x, L)p(x)\,dx \tag{3}$$

with $F(x, L)$ an appropriate Green's function [19, 37].

Taken together, Equations 1 to 3 provide us with the basis for determining threshold fracture configurations exclusively from surface force functions, $p(x)$. Unfortunately, the functional dependence $p(x)$ is not known *a priori*. To get around this problem we might attempt, given knowledge of the crack profile $y = y(x)$, to rewrite Equation 3 as an integral over the crack opening displacement y rather than the in-plane coordinate x. Then we could perhaps obtain an appropriate surface force function $p(y)$ by independent means, e.g. by direct measurements using the microbalance devices or from theoretical potential function calculations. However, the crack profile itself must be

dependent on $p(x)$, so decoupling the x-coordinate from the analysis is by no means straightforward. One has to evaluate a non-linear integral equation to obtain a self-consistent solution, which generally calls for numerical analysis. This is a common problem in fracture [38] and adhesion mechanics [39], and will be addressed in mathematical detail in our companion paper [23].

There are more serious difficulties, of a physical rather than mathematical nature, with the above description, namely the implication that at incremental departures from equilibrium the crack should extend (or recede) spontaneously. In reality, such departures take us into the realm of kinetic crack responses, i.e. on to a v–G curve. This calls for the introduction of activation barriers into the cohesive stress formalism, as for instance in the lattice trapping models of Thomson and co-workers [19, 40, 41]. Another constraint (which will become more evident in the next section) is the lack of any distinction between initial zero-velocity and subsequent healing–repropagation thresholds during crack loading–unloading–reloading cycles.

The implied existence of activation barriers in the surface force–separation formalism would appear to be consistent with the experimental observations of Israelachvili and co-workers [24, 33, 34] (Section 2.1). Accordingly, let us pursue their simplistic arguments based on elastic sphere representations of intervening molecular species between two separating, smooth, hard walls. Fig. 2 schematically illustrates the envisaged process. We plot energy–separation and force–separation diagrams for the surfaces with integral numbers of intervening molecular layers. Thus, Curve 0 corresponds to the separation of surfaces in the absence of any environmental interaction, i.e. to the primary cohesive forces, Curve 1 to separation with just a monolayer admitted to the interface, and so on. In constructing these diagrams it is implicit that the interacting molecules be allowed to adjust their positions between the walls so as to minimize the system free energy. For large separations the molecules are relatively free to make these adjustments, and so to attain their ordinary fluid state; but as the surfaces approach each other, toward single-layer coverage, this flexibility is clearly restricted. In short, there is an increasingly strong tendency to ordering of the constrained molecular layers as the interface narrows down. This accounts for the oscillations in the plots. According to this interpretation we would always expect a pronounced first subsidiary minimum (as indicated in Fig. 2), even in those non-ideal instances where "atomically rough" surfaces "wash out" the longer-range order [24].

Suppose then that we start with the system in its initial, primary equilibrium state and separate the surfaces to infinity. Consider first the special case where the environmental species are totally excluded from the interface, i.e. inert conditions. Under these conditions the system follows Curve 0 all the way, and the appropriate surface energy term is γ_0. Now take the case where the interface is connected to the environmental reservoir throughout the separation

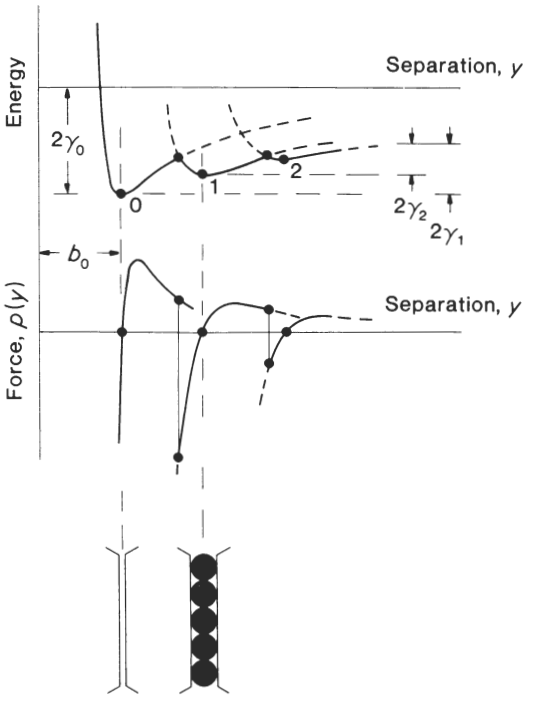

Figure 2 Energy–separation and force–separation diagrams for two (hard, smooth) surfaces with intervening interactive (elastic, spherical) molecular species. The position of the primary cohesive minimum along the y-axis is determined by the characteristic lattice spacing b_0 of the solid. Positions of the secondary minima are determined by the size of the intervening molecules. (In accordance with the coordinate system in Fig. 1, y and b_0 are half-spacings, measured relative to the crack symmetry plane.)

process. The system then follows the path of minimum energy, i.e. branching progressively from Curve 0 to Curves 1, 2, etc. Note that the external species, by virtue of their size, cannot even enter the interface until the walls have been separated by almost one molecular diameter. For this configuration the final energy level is reduced relative to its inert-state counterpart, by an amount equal to the adsorption energy. With the proviso that we can effect passage over the energy barriers in Fig. 2 reversibly, the appropriate surface energy becomes γ_1. This proviso concerning reversibility bears closer scrutiny, for there is the suggestion, through the presence of the potential function minima, that the system could exist in metastable equilibrium states. On attempting to bring two separated surfaces back to their primary equilibrium it might not be easy to surmount the final barriers: in other words the system may relax into a quasi-equilibrium state, with one or more molecular layers trapped at the interface. There is direct experimental evidence to support this trapped-layer notion: Bailey [5] measured the thicknesses of mica sheets before and after cleavage, using multiple-beam interferometry, and found a detectable increase of 0.6 nm (cf. water molecule diameter, 0.28 nm); Chan and Horn [42] performed mica–mica squeezing tests in organic liquids and found discrete, molecular-scale intervals in the approach distance against time data. If we were subsequently to re-separate the surfaces from, say, the first secondary minimum in Fig. 2, the apparent surface energy would be reduced still further, to γ_2.

With this description we postulate the v–G behaviour

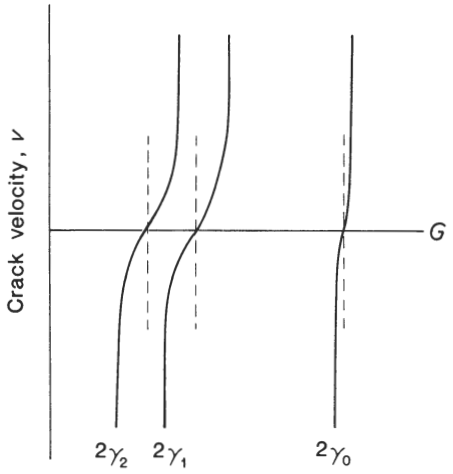

Figure 3 Schematic diagram of allowable v–G relations (after Rice [2]). The thermodynamic quiescent (zero velocity) points are determined by surface energy terms, as defined in Fig. 2. The $2\gamma_0$ curve corresponds to crack growth in environment-free conditions, the $2\gamma_1$ curve to growth with interactive species adsorbed and desorbed during forward and backward motion through virgin material, and the $2\gamma_2$ curve to similar growth but through a healed interface.

shown schematically in Fig. 3. The construction in this diagram is essentially the same as that described in the classic paper on thermodynamic fracture states by Rice [2]. The curve with threshold at $2\gamma_0$ corresponds to reversible crack growth under inert conditions; that at $2\gamma_1$ corresponds to reversible growth through virgin material in a reactive environment, with species adsorbed and desorbed during forward and backward motion; and that at $2\gamma_2$ corresponds to reversible growth in the same reactive environment through a healed interface with a trapped layer. For non-inert conditions, if desorption were to be inhibited as suggested above, the v–G response would transfer from the γ_1 to the γ_2 curve on the first unload to zero velocity, in accord with general observation. Let us note that there is nothing in our description to preclude negative γ terms; we simply indicate here that in such cases the thresholds correspond to imaginary states, such that the crack has non-zero velocity even at $G = 0$ [14].

3. Comparison with experimental observations in near-threshold region

Now let us examine some available data on mica, fused silica glass and sapphire in the light of the preceding crack-threshold considerations. We have indicated that systematic experimental studies of this kind are sparse. Here we shall describe some recent results from our own laboratories, and supplement these with others from the open literature. In all instances the data will be presented on crack velocity plots, to emphasize the kinetic as well as the equilibrium aspects of the phenomenon. The abscissa in these plots will be taken as G rather than the more conventional K_a so that the threshold points may be interpreted as appropriate surface energy quantities.

3.1. Mica

Cyclic fracture experiments were run on muscovite mica using a derivative of the wedge-loaded cantilever technique developed by others [4–8]. Crack motion was controlled by driving a stainless steel blade (thickness 0.2 mm), via a micrometer screw fixture, into the mica sheets (dimensions 15 mm × 10 mm × 0.05 mm). The entire specimen-drive fixture was mounted on to the stage of a microscope so that the crack interface could be continuously observed, in our setup by means of Fizeau fringes in reflected mercury light. Each loading cycle was effected in three abrupt stages, in a prescribed environment, as follows: (1) LOAD (L) by advancing the blade, to run the crack through virgin material; (2) UNLOAD (U) by partially withdrawing the blade, to allow the crack to retract; (3) RELOAD (R) by reinserting the blade almost (but never quite) to the first load position. The interfacial fringes allow us to distinguish true healing from mere closure; in the first case the entire fringe system translates with the crack tip, whereas in the second the fringe spacing simply expands. Ours is a stable crack configuration, i.e. G diminishing with crack length [43],* so that kinetic effects at any stage of the L–U–R sequence were always manifest as ever-diminishing velocities on approach to the new equilibrium state. With this arrangement more than one cycle could be run on a given sheet, thereby avoiding specimen-to-specimen scatter.

Fig. 4 shows the results of such cyclic tests on a single mica specimen, in air (relative humidity 55%) and water (distilled) environments. Here we have plotted the velocity ordinate in positive and negative logarithmic units (consistent with rate theory for reversible activation over barriers [45]). In these tests the time to run each half-cycle was about two hours. It is immediately clear that the slopes of the v–G curves are very steep (as compared, for instance, to silica glass, Section 3.2), indicative of strong threshold behaviour. Also, there are significant shifts along the G axis between the individual L, U, R branches of the loading cycle. This apparent hysteresis was noted by Bailey in her early experiments [5]. If in accordance with our discussion in Section 2.2 we associate the L branch with crack-plane separation from the primary minimum in the surface force function (Fig. 2), and likewise the U–R branches with (reversible) separation from the first secondary minimum, we may adopt the Rice scheme in Fig. 3 to estimate the surface energies γ_1 and γ_2. These energy terms correspond to the intersection points on the G axis, which must be determined by the somewhat uncertain process of interpolation (γ_2) or, even worse, extrapolation (γ_1).

From Fig. 4 we obtain for the healing–repropagation threshold in mica $\gamma_2 = 300\,\mathrm{mJ\,m^{-2}}$ in air and $50\,\mathrm{mJ\,m^{-2}}$ in water. These values are within a factor of two of those reported under similar conditions by Bailey and co-workers [5, 7]. In view of the uncertainties just mentioned in the data analysis, together with evidence of considerable specimen-to-specimen scatter in the work of those [5, 7] and other [8] authors, this

*In computing G we use beam theory [43], with the appropriate cleavage in-plane Young's modulus $E = 170\,\mathrm{GPa}$ [44].

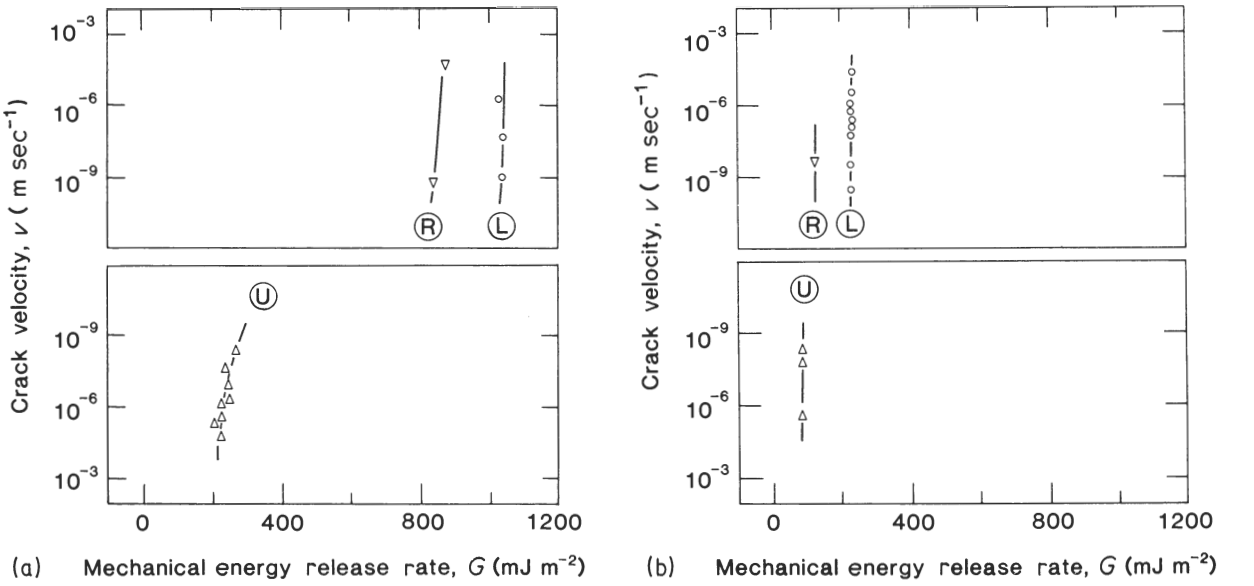

Figure 4 Crack velocity against G for mica in (a) air (55% relative humidity) and (b) water environments, for load–unload–reload (L–U–R) cycles. Curves are simple fits to data.

level of agreement is perhaps as much as we might expect. On the other hand, our value for air is well in excess of the so-called adhesion energies reported by Horn *et al.* [39] in resealed mica–mica pulloff tests, 40 to 80 mJ m^{-2}, in nitrogen gas (but after exposure to air). As pointed out by Horn *et al.*, the adhesion test is not without its own sources of uncertainties. With this apparent lack of consensus it would seem unwise to attach too much quantitative significance to the extrapolated γ_1 values for the initial loading threshold in Fig. 4, other than that they must be greater than their γ_2 counterparts. To compound the uncertainties concerning absolute energy determinations still further, we have noted additional complicating factors in our experiments, including "ageing" effects (e.g. hold time between half-cycles), "distance" effects (distance behind original virgin crack tip over which healing is allowed to occur), that can cause the v–G curves to shift about. Notwithstanding all these variations it is clear that we are dealing with non-trivial energies, of the order of tens or even hundreds of mJ m^{-2}, that are highly sensitive to the mica-environment surface chemistry.

3.2. Silica glass

Silicate glasses have long been adopted as the archetypal class of brittle material by the ceramics community. We have run cyclic fracture tests on fused silica microscope slides (75 mm × 25 mm × 1 mm) using a double torsion arrangement, analogous to the healing experiments of Stavrinidis and Holloway [11]. Our choice of pure silica rather than, say, soda–lime glass simply reflects a desire to work with "simple, well-defined" materials, although we have been running tests on other glass compositions. Cracks were started from a diamond scribe mark at one end, along the midplane, of the specimens. These were propagated into the centre region of the slide where constant G conditions† generally obtain [46]. Again, the entire specimen-support fixture was mounted on to a microscope stage to allow the subsequent steady-state forward or backward crack motion for any given applied load to be followed directly (transmitted light, crossed polarizers). The resultant stress birefringence associated with the moving crack tip field ensured that we were indeed observing true healing in our load cycling tests.

Results for L–U–R runs on one silica specimen in air (55% relative humidity) are shown in Fig. 5, using the same logarithmic plotting scheme as in Fig. 4. Crack reversal was effected immediately on completing the lowest velocity measurement for each half-cycle. These results agree to within 20% in G with those of Stavrinidis and Holloway [11] for similar test conditions (five minute hold time between half-cycles). It is interesting to note here that we see no sign of

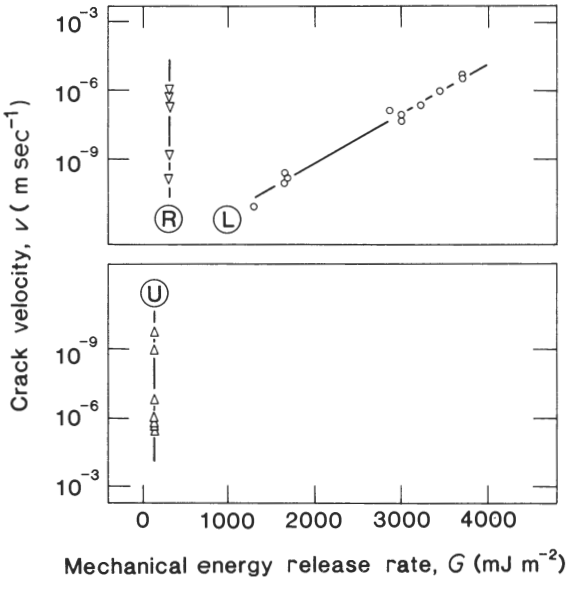

Figure 5 Crack velocity against G for silica glass in air (55% relative humidity), for L–U–R cycle. Curves are fits to data.

†G computed for $E = 70$ GPa.

a zero-velocity threshold on the L curve, but it is possible that we have not extended our data range to low enough velocities for this material. And of course the crack motion does reverse at lower G levels, indicating that such a threshold must exist, even if coincidentally with the healing–repropagation threshold. At this latter threshold, interpolation of the U–R data yields $\gamma_2 = 100\,\mathrm{mJ\,m^{-2}}$ for our silica–air system (i.e. comparable with the mica–air system). Michalske and Fuller [12] report a value $75\,\mathrm{mJ\,m^{-2}}$ from their healing studies.

It might appear from the above that glass represents a reasonably "clean", reproducible fracture system. Unfortunately, this is not so. As with mica we find evidence that the threshold energies can be greatly influenced by such factors as age time between half-cycles (see also [11]) and distance over which the crack is allowed to heal. Also, the water content of the environment can be important [12]; indeed, in liquid water it is difficult to get the crack to heal at all. The picture is more complicated in soda–lime and other mixed oxide glasses, where L-curve thresholds begin to appear; in these cases the influence of mobile cationic species in the bulk structure can become a dominant factor.

3.3. Sapphire

Sapphire is another material that has received attention from the fracture testing community, mainly because of its availability in single-crystal form. Wiederhorn [47] was the first to investigate the crack velocity behaviour of this material: he obtained v–K_a data for double-cantilever beam tests in air, showing steep slopes at low velocities (reminiscent of the steep slopes in the corresponding plots for mica, Fig. 4a). More recently, Michalske *et al.* [48] reported analogous results for tests in liquids, including water. Again the slopes were steep. We reproduce the latter results in Fig. 6.* Unfortunately, no data have yet been obtained for unloading–reloading sequences in sapphire (although we recall from Section 1 that sapphire does exhibit exhibit healing tendencies [13]).

A revealing comparison between the above sapphire–water crack velocity data and corresponding strength against stressing rate data from tests on specimens with controlled indentation flaws has recently been made by Cook [36]. By deconvoluting the latter data Cook was able to obtain an independent evaluation of the velocity curve; and one, moreover, that covered a far wider range of effective G values than before. His curve is included as the broken line in Fig. 6. Significantly, Cook was able to demonstrate the existence of a well-defined crack velocity threshold (clearly manifest as a fatigue limit in the strength data). Most interestingly, the data of Michalske *et al.* [48] in Fig. 6 lie in a region dominated by this threshold: it is only at velocities above $10^{-4}\,\mathrm{m\,sec^{-1}}$ that the curve begins to show the familiar strong G dependence. This would suggest that it is the kinetics associated with interfacial barriers (Section 2.2), and not with concerted chemical reactions, that should be used in analysing the results in Fig. 6.

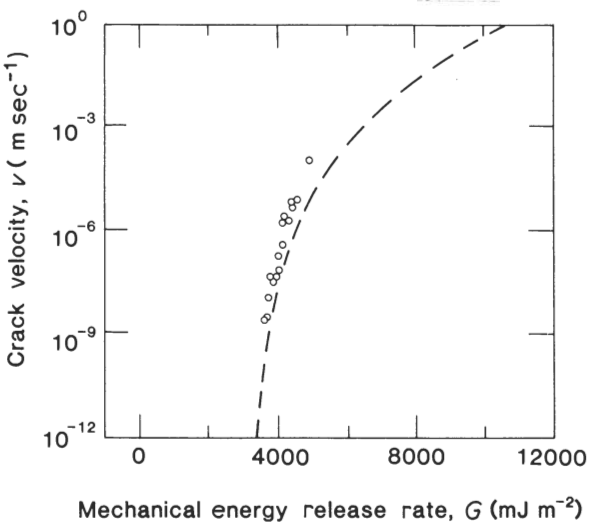

Figure 6 Crack velocity against G for sapphire in water, for loading through virgin material only (L). Data points represent direct velocity measurements (from [48]). Dashed curve represents deconvoluted v–G function from strength against stressing rate data (from [36]).

Given this interpretation, we may extrapolate the water data in Fig. 6 to zero velocity to obtain (approximately) $\gamma_1 = 2800\,\mathrm{mJ\,m^{-2}}$ [36]. Even allowing for the obvious uncertainty in this estimate (magnified no doubt by the considerable specimen-to-specimen scatter reported by the authors [48]), it is seen that this threshold is substantially higher than for mica or silica glass.

4. Atomistic modelling of crack interface structure

We have suggested that atomic-scale discreteness in the fundamental force–separation function can be a vital factor in determining fracture thresholds. We have implied that this discreteness is essential in accounting for the non-coincidence of L and U–R equilibrium states in loading–unloading–reloading cycles, and presumably for the attendant velocity curves on displacing the system away from these equilibrium states. In the present section we shall explore this element further, with more detailed attention to the geometrical constraints that relate to the accommodation of molecular species at crack interfaces in our specific structures. In keeping with the spirit of our constructions in Fig. 2 we shall retain the elastic sphere approximation as the basis for modelling these structures. This approximation, while admittedly limited in its power to represent real molecular systems, will nevertheless provide critical insights into the nature of the crack–environment interactions, insights not obtainable from the usual point-mass representations (and certainly not from strict continuum-based representations).

In the following subsections we shall begin with a consideration of the atomic structures of our model materials. As will become evident, the key to the stacking configurations in the structures lies largely

*Converted from K to G assuming isotropic elasticity, $E = 400\,\mathrm{GPa}$.

with the oxygen sublattices. Noting that the bonding which characterizes our materials is predominantly electrostatic,† we use ionic radii as follows [49]: O^{2-} 0.140 nm, K^+ 0.133 nm, Si^{4+} 0.041 nm, and Al^{3+} 0.050 nm; for H_2O, our principal environmental molecule, we take an averaged radius 0.14 nm.

Then we shall consider the distortions of these structures at cleavage crack interfaces in relation to the accommodation of the environmental molecules. In computing the interface profiles we resort to the simplistic Irwin displacement field solutions for cracks with traction-free walls in a linear isotropic continuum (see Fig. 1)

$$u_x = \left[\left(\frac{G}{2E}\right)\left(\frac{r_n}{2\pi}\right)\right]^{1/2} f_x(\theta_n, \nu) \quad (4a)$$

$$u_y = \left[\left(\frac{G}{2E}\right)\left(\frac{r_n}{2\pi}\right)\right]^{1/2} f_y(\theta_n, \nu) \quad (4b)$$

where $r_n = (x_n^2 + y_n^2)^{1/2}$, $\theta_n = \arctan(y_n/x_n)$, and the $f(\theta_n, \nu)$ are angular terms (evaluated separately for plane stress and plane strain) [43, 50]. Consistent with the elastic sphere picture we take the crack to be centred between adjacent atom (oxygen) planes, focusing our attention on the displacements of the atom centres located on these planes: that is, we evaluate Equations 4 at the discrete points $x_n = \pm na_0$, $y_n = \pm b_0$, with n an integer and a_0, b_0 characteristic "lattice" spacings. It will be noted that the approximation of traction-free walls here seems to fly in the face of our earlier assertions concerning the importance of surface force effects (Section 2). We justify this approximation solely on the grounds of expediency, to allow us to bring out a central point of our thesis without mathematical complexity, and discuss the consequences in more detail in Section 5. At the same time let us emphasize that our approach, by virtue of its attention to actual atom positions (as distinct from the positions of the hypothetical, infinitesimally narrow continuum slits usually treated by the fracture mechanics community) does incorporate the critical element of discreteness referred to above.

4.1. Mica

A point-mass diagram of the mica structure, $KAl_2(AlSi_3O_{10})OH_2$, is shown in Fig. 7 [51]. The essence of the structure is the arrangement of silica tetrahedra into double sheets, strongly linked at the inward-pointing tetrahedral vertices by aluminium ions (with hydroxyl groups incorporated into this internal structure). The double sheets are effectively bounded by oxygen layers, relatively weakly linked by Coulombic attractions to their neighbours by potassium ions. It is between these adjacent sheets that cleavage occurs. The corresponding hardsphere representation is given in Fig. 8, retaining only those atoms that participate in nearest-neighbour bonding at the cleavage plane. The oxygen layers are arranged in symmetrical arrays of hexagonal rings into which are recessed the coordinating potassium ions.

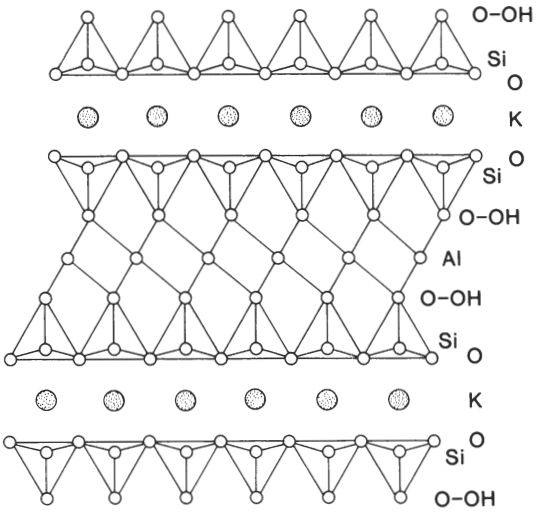

Figure 7 Point mass representation of mica structure, showing projection on to (1 0 0) plane (from [51]). Cleavage occurs along (0 0 1) plane, between potassium-coordinated oxygen layers.

The question may now be raised as to how this structure might accommodate penetrating water molecules at the cleavage plane. It soon becomes clear that the available interstices in the undistorted structure are not large enough for such penetration to occur without encountering significant diffusional energy barriers. It might be expected that water molecules could occupy median-point sites between any three potassium ions in Fig. 8a: indeed, there is no difficulty in accommodating such a fit laterally within the cleavage plane (Fig. 9a). However, in the orthogonal direction the separation between opposing oxygens that define the (six-fold coordination) interstitial sites is too small. To facilitate a fit in this direction we must separate the oxygen layers by some 50%, namely by 0.16 nm relative to the equilibrium spacing 0.34 nm, as in Fig. 9b. Note further that we have

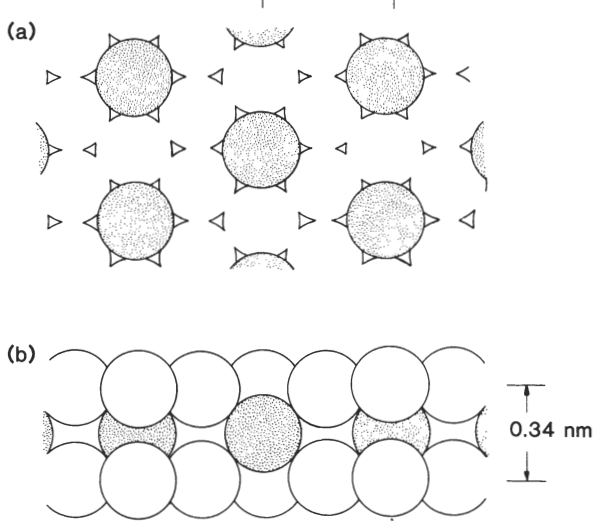

Figure 8 Hard-sphere model of mica structure, showing (a) projection and (b) profile views of cleavage plane. Potassium ions (shaded spheres) provide cohesion for the oxygen layers (open spheres).

†The degrees of ionicity for the important bonds, from Pauling's electronegativity tables [49], are 86% for K–O (mica), 51% for Si–O (silica) and 63% for Al–O (sapphire).

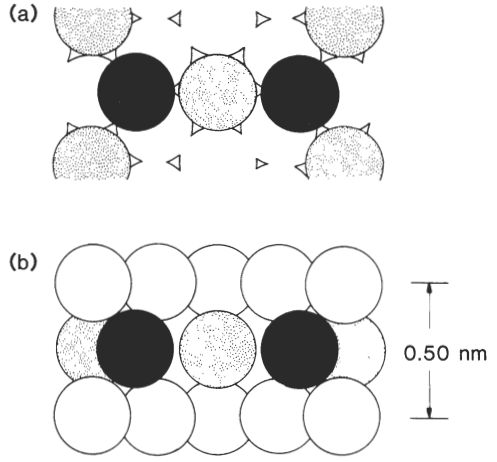

Figure 9 Accommodation of water molecules (dark spheres) at mica cleavage plane; (a) projection and (b) profile views. Whereas molecules can be accommodated laterally without strain between oxygens in (a), layers in (b) must first be separated by about 50%.

considered the most favourable accommodation sites here. Given that the configuration in Fig. 9b is attained, then the oxygen layers will need to be forced even further apart form the water molecules to pass from one such favoured site to the next.

Now let us turn to the corresponding crack configurations. Fig. 10 shows a sequence for growth into virgin material (L-curve) at increasing values of G, embracing the data range in Fig. 4.* In constructing these diagrams we show the potassium ions as alternately attached to the top and bottom surfaces, consistent with the requirements of charge neutrality. The water molecules we show located in the interstices of Fig. 8, back to the point along the interface at which the wall separation is sufficient to accommodate more than one molecular layer. At the largest G level represented in Fig. 10d it appears (within the limits of our earlier approximations embodied in Equation 4) that water molecules might have no difficulty in gaining access to the crack tip bonds. However, this G level lies well to the right of the data range in the experimental velocity plots of Fig. 4; it is the configurations of Figs 10a to c that are more properly representative of the near-threshold region of primary interest to us here. Thus as G diminishes the ensuing interfacial constriction will present an ever-increasing barrier to molecular penetration until, at a sufficiently low level, the "tip", as defined by the Irwin singularity at $x = 0$, $y = 0$, becomes totally inaccessible. Note that at the lowest G level in Fig. 10a the crack walls remain separated by less than one molecular dimension at up to tens of atom distances behind the Irwin tip.

If Fig. 10 represents the crack configurations for the initial loading stage, what are the corresponding representations for the subsequent unloading–reloading stages? Recall from Section 2.2 that we might expect one or more molecular water layers to become "trapped" at the interface during an opening–closing cycle. Fig. 11 compares appropriate crack profiles for growth through virgin and healed material, at the U–R

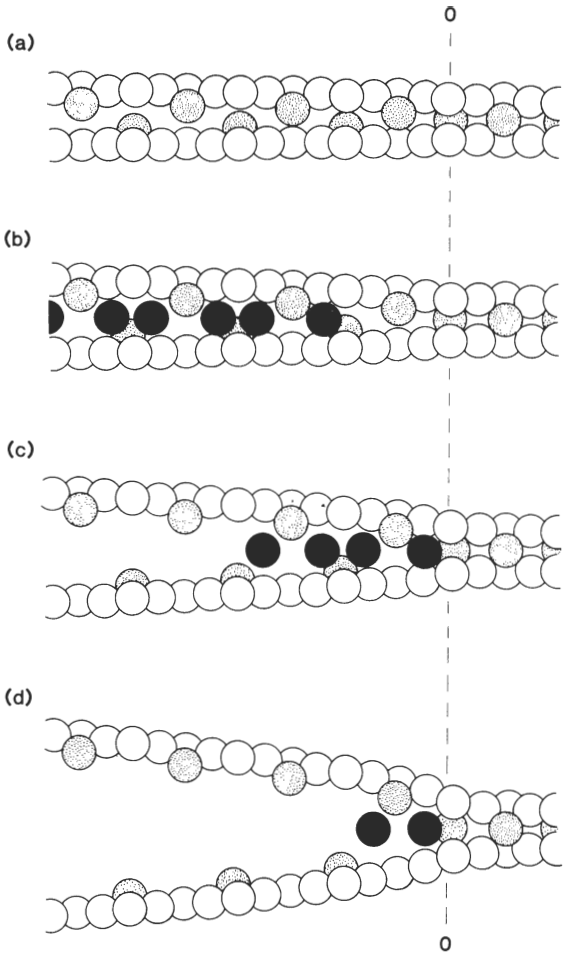

Figure 10 Irwin crack profiles for mica, at four G levels, for growth through virgin material (L-curve) in water: G = (a) 50, (b) 200, (c) 800, (d) 3000 mJ m^{-2}. Vertical dashed line 0–0 represents Irwin crack tip (i.e. origin $r = 0$ in Equation 4). Note severe geometrical restriction at interface in threshold (low G) regions.

threshold point for water, $G = 2\gamma_2 = 100\,\text{mJ m}^{-2}$ (Fig. 4b). In computing the profile in Fig. 11b we have made due allowance for the increase in latter spacing required to accommodate one layer of water molecules, i.e. as in Fig. 9b instead of Fig. 8b (equivalent to taking b_0 at the first secondary minimum rather than at the primary minimum in Fig. 2). If we concede water as deleterious to the cohesion then it is clear that the G level required to maintain "equilibrium" states in the U–R stages of the cyclic loading must be lower than that in the preceding L stage (consistent with a transition from $2\gamma_1$ to $2\gamma_2$, Fig. 2), as observed experimentally.

There are some possible variants on the simplistic model depicted in Figs 10 and 11 that warrant brief mention here. First, we have concerned ourselves only with water molecules in the environment. That other molecular species can make their presence felt is apparent from Fig. 4, by the way the v–G curves for air are displaced well to the right of those for water. It may be envisaged here that nitrogen molecules (say), even though relatively passive, could play an important role by occupying interstitial sites that would otherwise be taken up by water, effectively leading to

*For these calculations we use the out-of-plane Young's modulus $E = 60$ GPa [44] in Equation 4, to allow for the strong elastic anisotropy in mica.

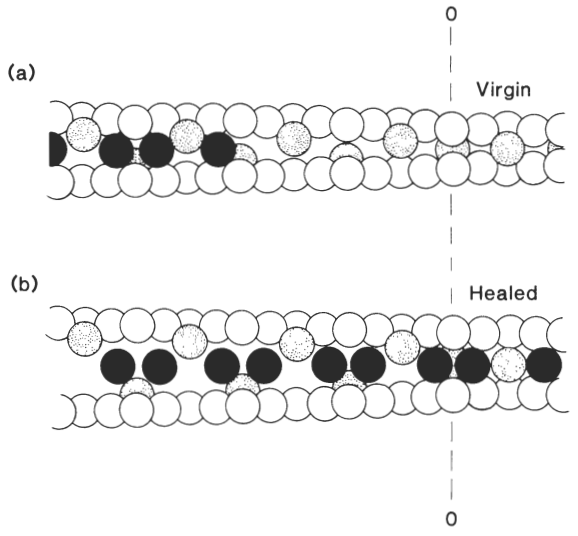

Figure 11 Irwin crack profiles for mica, comparing configurations for growth through (a) virgin and (b) healed interfaces. Constructed for G level at U–R threshold of $100\,\mathrm{mJ\,m^{-2}}$.

an increase in γ_2. Again, we have sketched only the single-layer configurations for the water molecules in our diagrams. In the less constrained regions further back from the crack tip, where the intruding molecules are more free to assume their usual fluid-like character, we might anticipate corrosion processes (including ion-exchange [52]) to become a factor (via their potential influence on the final energy states of the ruptured crack walls).

4.2. Silica

The structure of silica glass is not so well defined. However, the nearest-neighbour coordination of both silicon and oxygen ions is the same as that in any of the crystalline forms of SiO_2 [53]; it is simply the irregular stacking of the Si–O linkage tetrahedra that distinguishes the amorphous form. Consequently, we shall represent the structure by that of tridymite,

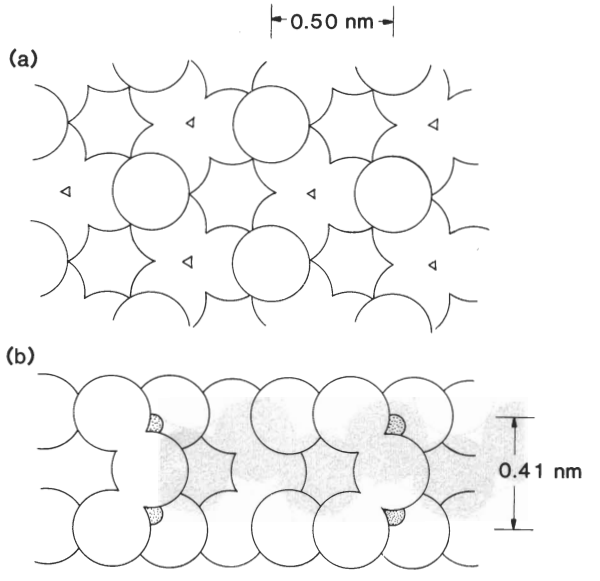

Figure 12 Hard-sphere model of SiO_2 (tridymite), showing network structure of silicon (small shaded spheres) and oxygen (large open spheres) (from [53]): (a) projection and (b) profile views of crack plane.

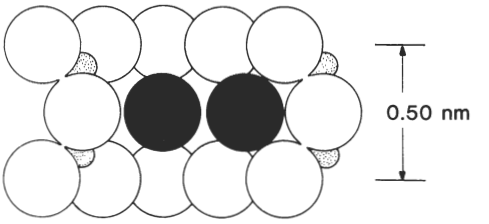

Figure 13 Accommodation of water molecules (dark spheres) at crack plane in silica, in profile view. Note two possible interstitial sites: left-hand site, into which a water molecule can fit without strain; right-hand site, which requires strain of about 20% for tight fit.

which has a density close to that of fused silica [53], with due acknowledgement that any quantitative aspects of our modelling will need to be regarded with caution.

Fig. 12 shows this structure in hard-sphere packing. Again it is a hexagonal ring array of oxygens (Fig. 12a) that provides the backbone of the structure. The relatively small silicon ions link these arrays into the tetrahedrally coordinated network. It is the Si–O–Si linkages (Fig. 12b) that must be severed for fracture to occur. Note that despite the seeming similarities in the oxygen substructure the atomic packing in this material is much less dense than it is in mica.

Consider now the accommodation of water molecules within this structure. We find this time that there are large "holes" into which the water can be fitted without elastic strain. These are the interstices located between adjacent oxygen hexagonal rings, as shown at left in Fig. 13. However, to get from one such interstice to the next the intruding molecules have to pass through the hexagonal rings, e.g. via the constrained (six fold coordination) site at right in Fig. 13: in this latter case a hard-sphere fit requires an opening strain of $0.05\,\mathrm{nm}/0.21\,\mathrm{nm} \simeq 20\%$ between oxygen layers. Thus although at first sight the silica structure does appear to be very "open" (even more so for an amorphous structure than Fig. 13 would indicate), it is manifest that the entry of external species must be limited by diffusion processes.

Fig. 14 illustrates these structural considerations

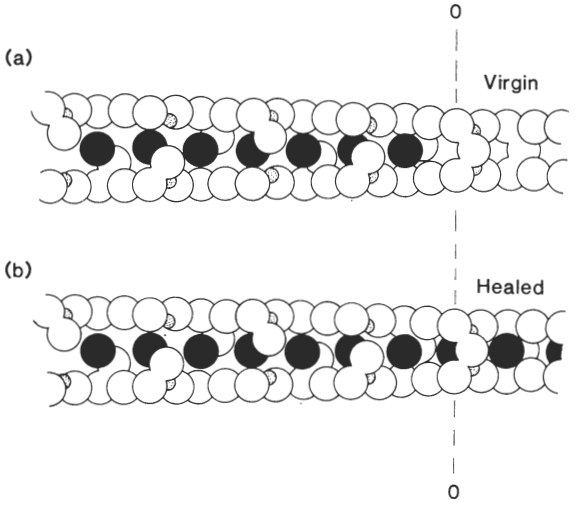

Figure 14 Irwin crack profiles for silica (tridymite) in water, at U–R threshold G level of $200\,\mathrm{mJ\,m^{-2}}$, for growth through (a) virgin material and (b) healed interface. 0–0 designates Irwin crack tip.

in relation to crack geometry, at the U–R threshold, $G = 2\gamma_2 = 200\,\mathrm{mJ\,m^{-2}}$ (see Fig. 5). We show configurations for growth in both virgin (Fig. 14a) and healed (Fig. 14b) interfaces. In the first of these configurations the water molecules are sketched in at the large holes, but only to the point at which the constrained six fold coordination sites form diffusion barriers. In the second configuration the water is assumed to occupy these open holes over the entire interface, corresponding to a healed crack with trapped layer. We acknowledge here that Fig. 14 may not present an entirely correct indication of the constrained interface structure, particularly if the water were to react chemically with the siloxane bonds at the crack walls (as is indeed assumed in the concerted-reaction theories of fracture [54]). However, such reaction products would serve only to add to the constraints experienced by diffusing species, so our conclusions here may be seen as conservative.

4.3. Sapphire

If the structure of silica is relatively open compared to mica, then that of sapphire, Al_2O_3, is just the opposite [51, 55]. The oxygen substructure is effectively a hexagonal close-packed array. Aluminium ions occupy two-thirds of the available octahedral interstitial sites between the oxygens. This incomplete occupancy preserves charge neutrality but, at the same time, leads to distortions from otherwise ideal lattice positions. We shall ignore the "puckering" of atomic layers that characterizes these distortions.

The cleavage of sapphire is somewhat ill-defined [56–58]. About all that can be stated with a degree of certainty is that (0 0 0 1) is not a preferred plane. Of the possibilities that have been studied we choose the $(\bar{1}\,0\,1\,0)$ plane, because of its crystallographic simplicity. This choice is not expected to limit the generality of our treatment. Fig. 15a is a hard-sphere projection on to this plane. (Note that the atomic layers are electrically nuetral here.) Adjacent atomic layers in this orientation are identical but laterally translated (by the partial lattice vector shown in the figure) so as to allow the oxygens to fit snugly in its close-packing arrangement (Fig. 15b). Clearly there is little room to accommodate foreign species.

Accordingly, the interposition of water molecules between the $(\bar{1}\,0\,1\,0)$ planes in Fig. 15 requires a high level of tensile strain. Thus, taking the most favourable interposition sites to be those previously occupied by the oxygens in the now-separated, adjacent layer (Fig. 16), we require a displacement/spacing ratio 0.32 nm/0.14 nm, i.e. a strain of some 230%. Clearly there are forbidding barriers to the entry of any external molecules.

Fig. 17 illustrates these configurations in relation to the crack threshold data for sapphire (Fig. 6). We show profiles for initial loading through virgin material at $G = 2500$ and $5500\,\mathrm{mJ\,m^{-2}}$ for a water environment. Again, the (Irwin-singularity) crack tip is seen to be inaccessible to the water molecules, and the threshold is explainable in terms of the influence of these molecules on the surface force interactions between the narrowly separated crack walls. It is difficult to conceive of any concerted reaction mechanism in this particular material at any G level within the data range of Fig. 6, contrary to customary interpretation [48].

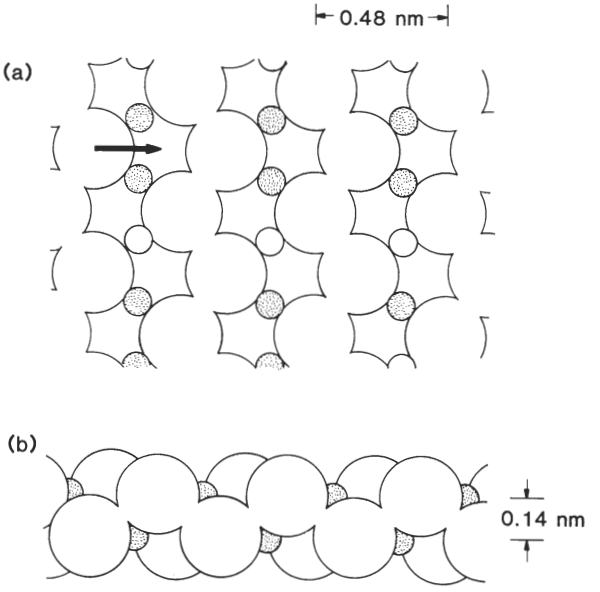

Figure 15 Hard-sphere model of sapphire structure (from [51, 55]): (a) projection and (b) profile views of $(\bar{1}\,0\,1\,0)$ cleavage plane. Aluminium ions (small spheres) occupy two-thirds of the available octahedral interstitial sites (occupation designated by shading) in otherwise close-packed oxygen (large open spheres) sublattice. Arrow in (a) indicates "lattice" translation vector between adjacent identical atom layers.

5. Discussion

We have investigated the underlying mechanisms of threshold behaviour in brittle materials retaining, perhaps beyond its admissible limits, the conventional Griffith–Barenblatt philosophy of equilibrium fracture. This approach has allowed us to formulate the problem in terms of thermodynamic surface energy states, in turn uniquely definable by fundamental surface force functions. The innovative element in our description is the introduction of molecular discreteness, consistent with the direct measurements of Israelachvili and co-workers, into these surface force functions. What emerges is a picture of the brittle crack that differs radically from the traditional open elliptical cavity: instead, we imagine a narrowly confined cohesive interface with atomic-scale barriers to the ingress (and egress) of reactive species. It is this picture, we believe, which holds the key to a complete fracture mechanics description of the v–G curve.

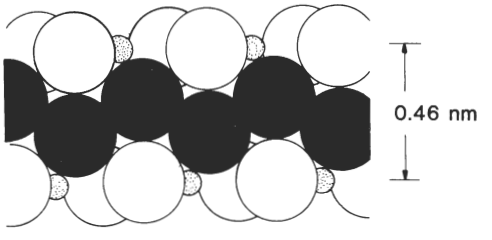

Figure 16 Accommodation of water molecules (dark spheres) at crack plane in sapphire, profile view. The wall separation needed to obtain a tight fit is especially large in this structure, corresponding to a strain of about 230%.

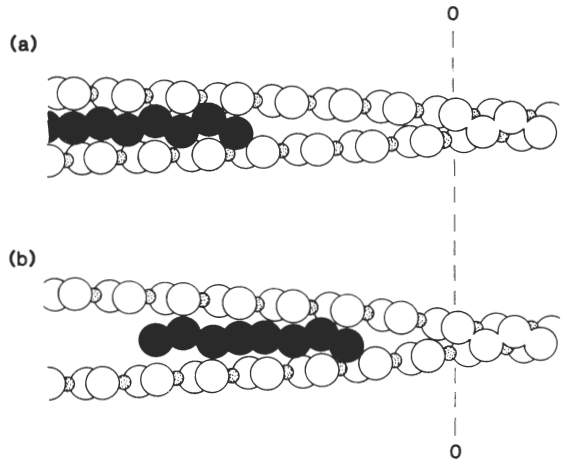

Figure 17 Irwin crack profiles in sapphire (virgin material) for L-curve in water: G = (a) 2500, (b) 5500 mJ m^{-2}. 0–0 designates Irwin crack tip.

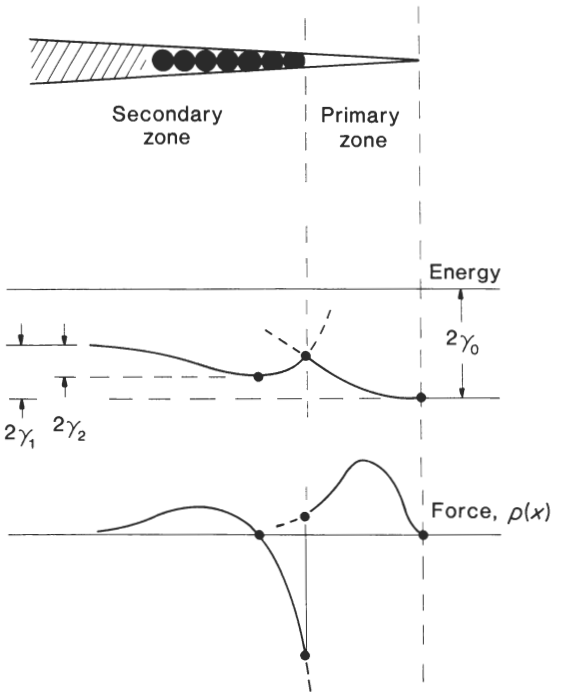

Figure 18 Schematic diagram of surface–surface interaction energy and force as function of crack-plane coordinate in cohesive zone. Note distinct separation into ("protected") primary zone, where the cohesion is determined exclusively by intrinsic interatomic bonding, and ("reactive") secondary zone, where all extrinsic interactions with intruding chemical species are confined.

Let us examine how such a description may be set up. Recall from Fig. 2 that the influence of active environmental molecules on surface–surface interactions is manifest as subsidiary minima in the force–separation or energy–separation functions. Of these minima it is the first that is most pronounced, by virtue of the stringent geometrical constraint associated with the interposition of a single layer of external molecules. We recall that atomic-scale roughness (Section 2.2), such as would be expected to characterize the true surface structure of silicate glasses, is expected to wash out higher-order minima. Accordingly, in the interest of simplicity we shall ignore these higher-order minima, noting that a more general description would serve only to introduce a possible multiplicity of metastable γ_2 states (Fig. 2). Now, we have alluded to the complexities which must attend any proper calculation of crack profiles with surface forces (Section 2.2); but, in so far as the wall–wall separation can be assumed to scale monotonically (albeit non-linearly) with distance behind the tip, we may take the (laterally inverted) plots in Fig. 2 to reflect the general x-dependence of the interaction functions. Fig. 18 is constructed in this spirit. Note that the lower plot corresponds to the cohesive stress function $p(x)$ needed to evaluate the integral in Equation 3.

Consider the implications of this modelling. The most striking feature is the clear distinction between primary and secondary cohesive zones. Remember that the boundary between these zones is determined by the configuration of maximum penetration for the given external species (see Figs 10, 14 and 17). The following points may be made concerning *equilibrium* states:

(i) The two zones, in so far as they remain spatially separated, are physically coupled via the elastic surrounds. Thermodynamically, energy expended in parting intrinsic bonds in the primary zone may be regained by adsorption of active species on to the now-active surfaces in the secondary zone. We might, however, expect dissipative "lattice trapping" losses [19] on transversing the hump in the energy curve in Fig. 18 to be greater for "narrower" cohesive zones (cf. "Peierls width" in dislocations). For an infinitely wide zone (zero dissipative loss) the energetics of surface formation are determined uniquely by the "end point" thermodynamic states in Fig. 18, i.e. as the difference between the final (ruptured) and initial (unruptured) states at $x = \infty$ and 0, respectively.

(ii) The concerted crack-tip reaction concept of environmentally assisted fracture may now be seen to be unnecessarily restrictive. Indeed, such a concept emerges as a special, limiting case of our present model; that in which the primary zone is of order one atomic spacing (strong "overlap" of primary and secondary energy functions in Fig. 18), such that the intruding molecules are able to interact with the 'still-intact" primary bonds. In view of our considerations in Section 4 we suggest that this picture is inapplicable in the threshold regions; sapphire, with its relatively dense packing (Fig. 17), serves to bring this point home most emphatically. Note that the surface force explanation of the L–U–R threshold phenomena, in particular the separation of the γ_1 and γ_2 states in Fig. 2, would survive even in this extreme: it is necessary only that the *secondary* cohesive zone should extend over more than a few atomic dimensions behind the tip.

(iii) Interactions with external species take place exclusively within the secondary zone. It is therefore implicit that (except in the limit of the concerted reaction) answers as to the nature of such interactions should be sought in surface chemistry (or more strictly, interface chemistry). This chemistry may involve, in addition to adsorption, such processes as ion-exchange and corrosion. Indeed, there are indications from our own observations of crack interfaces that the

occurrence of these additional processes may be the rule rather than the exception [59]. We reiterate that while reaction products may be spatially isolated from the tip they may nevertheless contribute strongly to the fracture mechanics, by virtue of their influence on the surface forces behind the crack tip. Note that if these processes continue to evolve with time at slower crack velocities the threshold conditions may themselves show some time dependence; our brief observations of ageing effects in Section 3.1 are consistent with this suggestion.

(iv) The conclusion that all chemical interactions with intruding species are confined to an isolated reaction zone provides us with a fundamental justification for our earlier assertions concerning the invariant structure of brittle crack tips [14–16, 20]. Our primary zone in Fig. 18 is "protected", by geometrical constraints, from exposure to these intruding species. As we have seen in Section 4, these constraints are severe in the threshold region for even open material structures (including silicate glasses) and small reactive molecules (including water). Thus contrary to those models which presume a change in the actual tip structure (e.g. blunting, as quantified by a tip radius [20]) we would assert that the criteria for crack extension remain uniquely expressible in terms of G (or some equivalent fracture mechanics parameter); the influence of environmental interactions is accommodated exclusively through variations in $p(x)$ in the internal cohesion function in Equation 3.

To this point our discussion has focussed on equilibrium aspects of the threshold phenomena. Can the scheme in Fig. 18 be used to account for the kinetics as well? Taking the argument made above in (i) at face value it might seem not: recall that for a wide cohesive zone the net barriers to translation of the crack, even for atomic-scale motions, are expected to be small; this means that the crack should extend or retract spontaneously for small displacements of G away from the appropriate 2γ level. It would therefore seem that Fig. 18 is oversimplistic in its representation of the crack–wall interactions. At this point we need to recognize that the energy–separation and force–separation functions in Fig. 2, from which Fig. 18 derives, take no account of *lateral* structure at the separation plane; the crack walls are effectively regarded as continuum-smooth. We have indicated repeatedly in the course of our hard-sphere modelling in Section 4 that this is not so. At low-G configurations the increasing incidence of constrained interstitial sites at the ever-narrowing interface must inhibit the transport of reactive species within the secondary cohesive zone.

Accordingly, it is suggested that non-equilibrium behaviour in the near-threshold regions must be controlled by diffusion kinetics. A theory of such kinetics, in particular the incorporation of a driving-force dependence into the velocity function, lies beyond the scope of this work. Nevertheless, we may make some pertinent remarks concerning this concept in relation to the **results in Section 3.** Most significantly, the "steepness" of the curves noted in Figs 4 to 6 may be attributed to the generally strong stress dependence associated with diffusion processes [60]. Thus, it is only at much higher G levels that the widely accepted "reaction-controlled" crack growth (so-called Region I)* processes can possibly begin to exert a dominant influence (as perhaps in the region of bend-over in the L-curve of Fig. 5, or in the broken curve representing Cook's data in Fig. 6), and even then the action of the surface forces may not be negligible (as reflected, for instance, in the slopes of the v–G curves [14, 36]). This conclusion has important implications as to the fracture mechanics basis of lifetime prediction in ceramics design, particularly in the context of fatigue limits.

A further point of interest concerning the above interpretation is the inevitable conclusion that the surface energy terms must exert a controlling influence in the fracture response, even in the kinetic regions. This conclusion has its roots in irreversible thermodynamics [2], whereby it can be shown that it is the quantity $g = G - 2\gamma$ (not G) that constitutes the true driving force for non-equilibrium fracture processes [45, 61]. It is interesting to note that a similar philosophy concerning true driving forces has been developed in adhesion mechanics [61]. Thus the threshold configurations may be properly regarded as critical reference states for the entire v–G response.

Let us now return to the expedient approximation made in Section 4 concerning the crack profile calculations. It will be recalled that we treated the crack surfaces as though they were free of any imposed stresses, in apparent contradiction to the very spirit of our thesis concerning the critical role of surface forces in the fracture process. This representation is clearly oversimplistic. As an illustration, consider the calculated profiles for mica in Fig. 10. The displacements of the crack walls in the near-tip region, up to the point of maximum penetration of water molecules, are too small (less than an atomic diameter) to take the surfaces beyond the range of primary cohesive interactions. These interactions will tend to pull the crack closed, effectively moving the "tip" to the left. Opposing this closure will be the water molecules, by exerting an increasingly effective wedging action on the walls (i.e. by enhancing the repulsive force evident at the primary–secondary zone boundary in Fig. 18). Thus we must expect the requirement of force equilibrium at each crack-wall atom site in the combined presence of external G and internal γ to give rise to extremely complex displacement fields in the near-tip region. It is in the context that it becomes necessary to resort to numerical solutions of the self-consistent problem [23]. Nevertheless, because of the mutually opposing influences in the tip region it may be argued that our crack profiles contain all the qualitative features of the molecular scaling and interface structural elements needed to account for activated crack growth.

To summarize, we have developed a basis for modelling brittle fracture in terms of fundamental

*In keeping with this terminology we might designate our diffusion-controlled region as "Region 0".

surface forces. In particular, we have given an account of threshold behaviour, with the Griffith–Barenblatt equilibrium mechanics appropriately modified to accommodate an essential element of atomic-scale discreteness in these forces. This link opens up the novel prospect of making *a priori* predictions of fracture behaviour, using direct surface force measurements (e.g. as in Israelachvili's apparatus). Conversely, it suggests that threshold crack configurations might be useful as a means for obtaining information on the forces themselves, especially for those materials not amenable to direct measurement (which at present means practically any material other than mica). At the same time, it should be recognized that many important questions remain to be answered. We have made no strong attempt to speculate as to the underlying nature of the interaction between crack walls and interposed species. Do the subsidiary minima in Fig. 2 arise because of a "dielectric screening" of the primary bonds [62, 63] or because of some additional "structural" forces [24]? We have noted that our materials are largely ionic (Section 4). What is the role of bond type? Again, in introducing discreteness into the description we have represented our cracks as "lattice" planes separated by atomic dimensions. How can we reconcile this representation with that of the infinitesimally narrow slit (Fig. 1) used as the basis for writing down the linear stress intensity factor relations in Equations 2 and 3? The very existence of a stress intensity factor implies a well-defined, if singular, crack tip. How does this notion of a crack tip fit with the absence of any abrupt $y(x)$ crack-plane discontinuities whatsoever in the configurations of Figs 10, 14 and 17? Notwithstanding these as-yet unresolved issues, the present study should serve as a useful starting point for a proper analysis of equilibrium and kinetic fracture in brittle solids at the fundamental level.

Acknowledgements

The authors wish to thank Charles Alpert for assistance in computing the crack profiles in Figs 10, 11, 14 and 17, Donna Johnston for painstakingly producing the finished drawings, and Deborah Heuckeroth for obtaining the data in Fig. 4. Funding for this work was provided by the US Office of Naval Research, Metallurgy and Ceramics Program.

References

1. B. R. LAWN and T. R. WILSHAW, "Fracture of Brittle Solids" (Cambridge University Press, London, 1975) Ch. 2.
2. J. R. RICE, *J. Mech. Phys. Solids* **26** (1977) 61.
3. E. D. CASE, J. R. SMYTH and O. HUNTER, in "Fracture Mechanics of Ceramics", Vol. 5, edited by R. C. Bradt, A. G. Evans, D. P. H. Hasselman and F. F. Lange (Plenum, New York, 1983) p. 507.
4. J. W. OBREIMOFF, *Proc. R. Soc.* **A127** (1930) 290.
5. A. I. BAILEY, *J. Appl. Phys.* **32** (1961) 1407.
6. P. J. BRYANT, L. H. TAYLOR and P. L. GUTSHALL, in "Transactions of Tenth National Vacuum Symposium" (Macmillan, New York, 1963) p. 21.
7. A. I. BAILEY and S. M. KAY, *Proc. R. Soc.* **A301** (1967) 47.
8. R. B. LEONESIO, *J. Amer. Ceram. Soc.* **55** (1972) 437.
9. S. M. WIEDERHORN and P. R. TOWNSEND, *ibid.* **53** (1970) 99.
10. G. L. CHEESEMAN and B. R. LAWN, *Phys. Status Solidi* **3** (1970) 951.
11. B. STAVRINIDIS and D. G. HOLLOWAY, *Phys. Chem. Glasses* **24** (1983) 19.
12. T. A. MICHALSKE and E. R. FULLER, *J. Amer. Ceram. Soc.* **11** (1985) 586.
13. B. J. HOCKEY and B. R. LAWN, *J. Mater. Sci.* **10** (1975) 1275.
14. B. R. LAWN, *Appl. Phys. Lett.* **47** (1985) 809.
15. D. R. CLARKE, B. R. LAWN and D. H. ROACH, in "Fracture Mechanics of Ceramics", Vol. 8, edited by R. C. Bradt, A. G. Evans, D. P. H. Hasselman and F. F. Lange (Plenum, New York, 1986) p. 341.
16. D. H. ROACH, D. M. HEUCKEROTH and B. R. LAWN, *J. Colloid Interface Sci.* **114** (1986) 292.
17. T. A. MICHALSKE, in "Fracture Mechanics of Ceramics", Vol. 5, edited by R. C. Bradt, A. G. Evans, D. P. H. Hasselman and F. F. Lange (Plenum, New York, 1983) p. 277.
18. R. J. CHARLES and W. B. HILLIG, in "Symposium sur la Resistance Mechanique du Verre et les Moyuns de L'Ameliorer" (Union Sciences Continentale du Verre, Charleroi, Belgium, 1962) p. 511.
19. R. M. THOMSON, *Ann. Rev. Mater. Sci.* **3** (1973) 31.
20. B. R. LAWN, A. C. GONZALEZ and K. JAKUS, *J. Amer. Ceram. Soc.* **68** (1985) 25.
21. D. H. ROACH and A. R. COOPER, *ibid.* **68** (1985) 632.
22. G. I. BARENBLATT, *Adv. Appl. Math.* **7** (1962) 55.
23. R. M. THOMSON and B. R. LAWN, in preparation.
24. J. N. ISRAELACHVILI, "Intermolecular and Surface Forces" (Academic, London, 1985).
25. A. W. ADAMSON, "Physical Chemistry of Surfaces" (Wiley, New York, 1982).
26. B. V. DERYAGUIN and L. LANDAU, *Acta Phys. Chem. USSR* **14** (1948) 633.
27. E. J. W. VERWEY and T. Th. G. OVERBEEK, "Theory of the Stability of Lyophobic Colloids" (Elsevier, Amsterdam, 1948).
28. D. TABOR and R. H. S. WINTERTON, *Proc. R. Soc.* **A312** (1969) 435.
29. J. N. ISRAELACHVILI and D. TABOR, *ibid.* **A331** (1972) 19.
30. J. N. ISRAELACHVILI and G. E. ADAMS, *J. Chem. Soc., Faraday Trans. I* **74** (1979) 975.
31. R. G. HORN and J. N. ISRAELACHVILI, *Chem. Phys. Lett.* **71** (1980) 192.
32. J. N. ISRAELACHVILI, *Phil. Mag.* **A43** (1981) 753.
33. H. K. CHRISTENSON, R. G. HORN and J. N. ISRAELACHVILI, *J. Colloid Interface Sci.* **88** (1982) 79.
34. R. G. HORN and J. N. ISRAELACHVILI, *J. Chem. Phys.* **75** (1981) 1400.
35. E. OROWAN, *Nature* **154** (1944) 341.
36. R. F. COOK, *J. Mater. Res.* in press.
37. R. M. THOMSON, in "Solid State Physics", edited by H. Ehrenrich and D. Turnbull (Academic, New York, 1986) vol. 39, p. 1.
38. D. B. MARSHALL, B. N. COX and A. G. EVANS, *Acta Metall.* **33** (1985) 2013.
39. R. G. HORN, J. N. ISRAELACHVILI and F. PRIBAC, *J. Colloid Interface Sci.* in press.
40. E. R. FULLER and R. M. THOMSON, in "Fracture Mechanics of Ceramics", Vol. 3, edited by R. C. Bradt, D. P. H. Hasselman and F. F. Lange (Plenum, New York, 1978) p. 507.
41. E. R. FULLER, B. R. LAWN and R. M. THOMSON, *Acta Metall.* **28** (1980) 1407.
42. D. Y. C. CHAN and R. G. HORN, *J. Chem. Phys.* **83** (1985) 5311.
43. B. R. LAWN and T. R. WILSHAW, "Fracture of Brittle Solids" (Cambridge University Press, London, 1975) Chs. 1, 3.
44. M. T. VAUGHAN and S. GUGGENHEIM, *J. Geophys. Res.* **91** (1986) 4647.
45. B. R. LAWN and T. R. WILSHAW, "Fracture of Brittle Solids" (Cambridge University Press, London, 1975) Ch. 8.
46. B. J. PLETKA, E. R. FULLER and B. G. KOEPKE, in

"Fracture Mechanics Applied to Brittle Materials", edited by S. W. Freiman (ASTM Special Technical Publication 678) (American Society for Testing and Materials, Philadelphia, 1978) p. 19.
47. S. M. WIEDERHORN, *Int. J. Fract. Mech.* **4** (1968) 171.
48. T. A. MICHALSKE, B. C. BUNKER and S. W. FREIMAN, *J. Amer. Ceram. Soc.* **69** (1986) 721.
49. L. PAULING, "The Nature of the Chemical Bond" (Cornell University Press, Ithaca, 1960).
50. G. R. IRWIN, in "Handbuch der Physik", Vol. 6 (Springer, Berlin, 1958) p. 551.
51. W. L. BRAGG and G. F. CLARINGBULL, "Crystal Structures of Minerals", Vol. IV (Bell, London, 1965) Chs. 6, 8.
52. P. M. CLAESSON, P. HERDER, P. STENIUS, J. C. ERIKSON and R. M. PASHLEY, *J. Colloid Interface Sci.* **109** (1986) 31.
53. R. K. ILER, "The Colloidal Chemistry of Silica and Silicates" (Cornell University Press, Ithaca, 1955).
54. T. A. MICHALSKE and S. W. FREIMAN, *Nature* **295** (1982) 511.
55. M. L. KRONBERG, *Acta Metall.* **5** (1957) 507.
56. S. M. WIEDERHORN, *J. Amer. Ceram. Soc.* **52** (1969) 485.
57. S. M. WIEDERHORN, B. J. HOCKEY and D. E. ROBERTS, *Phil. Mag.* **28** (1973) 783.
58. R. CANNON, in "Advances in Ceramics", Vol. 10, edited by W. D. Kingery (American Ceramic Society, Columbus, 1984) p. 818.
59. D. H. ROACH, S. LATHABAI and B. R. LAWN, *J. Amer. Ceram. Soc.*, in press.
60. R. M. BARRER, "Zeolites and Clay Minerals as Sorbents and Molecular Sieves" (Academic, London, 1978).
61. D. MAUGIS, *J. Mater. Sci.* **20** (1985) 3041.
62. G. L. GAINES and D. TABOR, *Nature* **178** (1956) 1304.
63. A. I. BAILEY and H. DANIELS, *J. Phys. Chem.* **77** (1973) 501.

Received 6 January
and accepted 4 March 1987